Improvement of Raman lidar algorithm for quantifying aerosol extinction

Felicita Russo[a], David Whiteman[b], Belay Demoz[b], Raymond Hoff[a]

[a] University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250
[b] NASA Goddard Space Flight Center, Greenbelt, MD 20771

e-mail: rfeli@umbc.edu

Popular summary:

Aerosols are particles of different composition and origin and influence the formation of clouds which are important in atmospheric radiative balance. At the present there is high uncertainty on the effect of aerosols on climate and this is mainly due to the fact that aerosol presence in the atmosphere can be highly variable in space and time. Monitoring of the aerosols in the atmosphere is necessary to better understanding many of these uncertainties. A lidar (an instrument that uses light to detect the extent of atmospheric aerosol loading) can be particularly useful to monitor aerosols in the atmosphere since it is capable to record the scattered intensity as a function of altitude from molecules and aerosols. One lidar method (the Raman lidar) makes use of the different wavelength changes that occur when light interacts with the varying chemistry and structure of atmospheric aerosols.

One quantity that is indicative of aerosol presence is the aerosol extinction which quantifies the amount of attenuation (removal of photons), due to scattering, that light undergoes when propagating in the atmosphere. It can be directly measured with a Raman lidar using the wavelength dependence of the received signal. In order to calculate aerosol extinction from Raman scattering data it is necessary to evaluate the rate of change (derivative) of a Raman signal with respect to altitude. Since derivatives are defined for continuous functions, they cannot be performed directly on the experimental data which are not continuous. The most popular technique to find the functional behavior of experimental data is the least-square fit. This procedure allows finding a polynomial function which better approximate the experimental data. The typical

approach in the lidar community is to make an *a priori* assumption about the functional behavior of the data in order to calculate the derivative. It has been shown in previous work that the use of the chi-square technique to determine the most likely functional behavior of the data prior to actually calculating the derivative eliminates the need for making *a priori* assumptions. We note that the *a priori* choice of a model itself can lead to larger uncertainties as compared to the method that is validated here.

In this manuscript, the chi-square technique that determines the most likely functional behavior is validated through numerical simulation and by application to a large body of Raman lidar measurements. In general, we show that the chi-square approach to evaluate aerosol extinction yields lower extinction uncertainty than the traditional technique. We also use the technique to study the feasibility of developing a general characterization of the extinction uncertainty that could permit the uncertainty in Raman lidar aerosol extinction measurements to be estimated accurately without the use of the chi-square technique.

Improvement of Raman lidar algorithm for quantifying aerosol extinction

Felicita Russo[a],David Whiteman[b],Belay Demoz[b], Raymond Hoff[a]

a University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250

b NASA Goddard Space Flight Center, Greenbelt, MD 20771

e-mail: rfeli@umbc.edu

We report on the improvement of a Raman lidar algorithm for calculating aerosol extinction. In order to calculate aerosol extinction from Raman lidar data it is necessary to evaluate the derivative of a molecular Raman signal with respect to range. The typical approach taken in the lidar community is to make an *a priori* assumption about the functional behavior of the data in order to calculate the derivative. It has previously been shown that the use of the chi-square technique to determine the most likely functional behavior of the data prior to actually calculating the derivative eliminates the need for making *a priori* assumptions. Here that technique is validated through numerical simulation and by application to a large body of Raman lidar measurements. In general, we show that the chi-square approach to evaluating extinction yields lower extinction uncertainty than traditional techniques. We also use the technique to study the feasibility of developing a general characterization of the extinction uncertainty that could permit the uncertainty in Raman lidar aerosol extinction measurements to be estimated accurately without the need of the chi-square technique.

OCIS codes: 010.1110, 280.0280, 290.2200, 010.3640

1. Introduction

The aerosol extinction coefficient is an important quantity in the study of the influence of aerosols on climate. At present there is a large uncertainty in the globally averaged forcing due to the aerosol indirect effect,[1] which is estimated to be between 0 and -2 Wm^{-2}. This uncertainty is related mainly to a poor understanding of cloud microphysics together with the fact that aerosols have very inhomogeneous distributions in the atmosphere that can significantly change with time. One of the important capabilities of Raman Lidar systems is the ability to retrieve range-resolved profiles of aerosol extinction.[2] This, together with the possibility of long-term monitoring of aerosols can help improve the current knowledge of the influence of aerosols on climate. EARLINET (European Aerosol Research LIdar NETwork)[3] is a network of lidars on a continental scale provides useful information to satisfy these needs. Raman lidar systems involved in EARLINET are located in 22[3] different geographic locations over the European continent. The lidar stations in EARLINET use a number of different techniques to calculate extinction. Particular attention has been given to the relative precision of the calculation of the aerosol extinction from each lidar station. In order to validate all the algorithms extinction comparison efforts have been executed.[4] As we will show, in comparison efforts such as this one the use of different algorithms can lead to large differences in the resulting extinction uncertainties. Moreover the *a priori* choice of the model itself can result in larger uncertainties as compared to the method that is validated here. We will refer to the frequently used technique of evaluating the aerosol extinction using a linear least squares fit[4] as the "traditional technique". We will show that, for the cases studied here, the linear model upon which the traditional technique is based is the least probable model to fit Raman lidar data for the purposes of evaluating aerosol extinction, and consequently gives on average a larger uncertainty than the technique described here.

The paper is organized in the following way: in section 2 the equations to be used and the chi-square technique will be described together with the findings from previous related work. In section 3 the results of simulations done with different aerosol extinction profiles will be presented. Section 4 will present the results obtained with some experimental data from the

NASA/Goddard Space Flight Center SRL (Scanning Raman Lidar) during two measurement campaigns in which it was deployed at the DOE (Department Of Energy)/ARM (Atmospheric Radiation Measurement) site in Lamont, Oklahoma. Section 5 will present the implications of the results of this work on the measurements of quantities important to the study of the aerosol effects on the climate and finally section 6 will provide a summary of the study with conclusions.

2. The Equations

The aerosol extinction coefficient is usually calculated from the Raman nitrogen signal with the equation:[2]

$$\alpha(\lambda_L, z) = \frac{\frac{d}{dz}\ln(\frac{N_N(z)}{z^2 P(\lambda_N, z)}) - \alpha_{mol}(\lambda_L, z) - \alpha_{mol}(\lambda_N, z)}{1 + (\frac{\lambda_L}{\lambda_N})^k} \tag{1}$$

where $\alpha_{aer}(\lambda_L, z)$ is the aerosol extinction at the wavelength of the laser λ_L, λ_N is the wavelength of the molecular nitrogen channel, $\alpha_{mol}(\lambda_L, z)$ and $\alpha_{aer}(\lambda_N, z)$ are the molecular extinction respectively, at the laser wavelength (λ_L) and at the nitrogen Raman wavelength (λ_N), $N_N(z)$ is the number density of the molecular nitrogen and $z^2 P(\lambda_N, z)$ is the range-square-corrected nitrogen signal. Here a scaling has been assumed for the aerosol extinction with wavelength denoted by the Angstrom coefficient $k(z)$ that is, in general, a function of the altitude, since it is a function of the kind of aerosol present. For this study, the Angstrom coefficient is put equal to one since the effort here will focus on the evaluation of the derivative.

In order to calculate the extinction it is necessary to calculate a derivative (Eq. (1)), a quantity that is defined for continuous functions. Here the argument of the derivative is not a continuous function instead it is lidar signal strength in discrete range cells, so the derivative must be calculated numerically. Different methods that are traditionally used for this calculation are sliding averages, as well as other smoothing techniques, and sliding linear fit.[4]

It has been shown[5] that the *a priori* assumption, represented by the choice of one of the

methods traditionally used, before executing the derivative is not necessary and that in fact it goes against the rules of statistics. The chi squared test will therefore be used here as a tool for avoiding the *a priori* assumption, since it allows to determine from the data the model that is most likely to represent the parent population of the data.

Before proceeding further, some considerations have to be made about the term containing the derivative in Eq. (1). This factor can be written as:

$$\frac{d}{dz}\ln\left(\frac{N_N(z)}{z^2P(\lambda_N,z)}\right) = \frac{1}{N_N(z)}\frac{dN_N(z)}{dz} - \frac{1}{z^2P(\lambda_N,z)}\frac{d}{dz}(z^2P(\lambda_N,z)) \tag{2}$$

The results presented in this paper are entirely obtained expressing the derivative as in Eq. (2). Since the number of accumulated counts required in a given range cell in order for the Raman lidar signal to be useful for deriving aerosol extinction is much greater than 10, we can assume that the statistics pertaining to each range cell is essentially Gaussian. However, as has been described before,[5] the argument of the derivative in equation 1 does not follow a Gaussian distribution. In fact the ratio of two Gaussian variables (as $N_N(z)$ and $z^2P(\lambda_N,z)$) in general does not possess Gaussian statistics and in the case of quantities of interest to the lidar community, the statistics of the ratio $N_N(z)/\ z^2P(\lambda_N,z)$ is at best approximately Gaussian.[6] The natural logarithm modifies the statistics further and statistical tests executed on simulated data, not shown in this paper, indicate that the argument of the derivative in Eq. (1) does not possess a Gaussian statistics in accordance with theory. Therefore it is not possible to meaningfully apply the chi-square test to the lidar data unless the aerosol extinction isevaluated using Eq. (2). Since such calculation is straightforward we suggest that this expression be used to calculate aerosol extinction instead of Eq.(1).

The choice of the most probable model is dependent on the chi square probability. The cumulative probability, which is defined as the integral of the probability distribution function, is used to determine how likely a given fit is and represents the probability of obtaining a chi-squared larger than that obtained in the fitting exercise. The measured chi square will be reasonably close to its expectation value as long as its cumulative probability is reasonably close to 0.5.[7] In our interpretation of this test among the set of models chosen, that in

this case are first to third order polynomials, the model which has a cumulative probability closest to 0.5 is considered the most probable.

Previous applications of the chi-square technique,[5] using just the linear and quadratic models, showed that differences in extinction as large as 10% were made when using this technique as opposed to the traditional technique. Moreover differences up to ± 40% were present among the corresponding uncertainties.

3. The Simulation

In order to show how the algorithm works a simulation was performed. A numerical model[8] was used to simulate Raman nitrogen signals as would be measured by a lidar system with the same operational characteristics as the NASA/GSFC Scanning Raman Lidar (SRL) during 1996-1997.[9,10] During this period, the SRL participated in three field campaigns at the DOE (Department Of Energy) ARM (Atmospheric Radiation Measurements) site in Lamont, Oklahoma. The experimental data that will be used in this paper are from WVIOP (Water Vapor Intensive Operation Period) held from 10 to 30 Sept 1996 and WVIOP2 held from Sept 15 to Oct 5 1997.[11] Each of these campaigns focused on quantifying accuracies and determining limitations of atmospheric water vapor measurement technologies. To reach this goal, measurements of atmospheric water vapor from different co-located sensors, including the SRL, were performed. Besides water vapor measurements, aerosol backscatter and extinction measurements were also acquired by the SRL during the three campaigns but were not previously analyzed. In section 4 will study these data using the chi-square technique described here.

The parameters that were used in the numerical model are reported in table 1. They were chosen to simulate the nighttime measurement characteristics of the SRL and were validated by selecting efficiency terms that force the simulation to match actual measurements. Differences between simulated and experimental data were reduced to less than 2.5% within the overlap region, decreasing with altitude.

In order to test the method in different conditions of extinction and signal uncertainty, two synthetic profiles of extinction were created and are shown in figure 1. One represents a

constant aerosol loading in the boundary layer (Ext1) and the other simulates a lofted aerosol layer (Ext2). To calculate the derivative of the nitrogen signal as a function of altitude, which is required to calculate the extinction profile from a nitrogen signal profile, a window of five points was selected. The resolution of the lidar signal is 75m and a window of five points corresponds to a cell 375m wide. The five-point window slides one point along the lidar profile for each successive determination of the extinction. A least square fit is performed on the five points in each cell. For each cell one value of the extinction and the corresponding uncertainty are calculated from the fit parameters. In the cases discussed in this paper, 3 orders of polynomials are considered and 3 different least square fits (linear, quadratic and cubic) are performed on each group of five points resulting in 3 extinction profiles with corresponding chi-square probability. The chi-square test is finally used to choose the most probable model at each altitude thus creating a single extinction profile as a function of the altitude. The sliding window creates a profile of extinction and uncertainty with the same spatial resolution as the lidar original data.

In this simulation the statistical performance of the algorithm was investigated by using an ensemble of 200 synthetic lidar profiles of Raman nitrogen. With all other parameters remaining constant, different averaging times were simulated so that a range of random uncertainties in the lidar signal could be investigated. Since, as it will be shown, the behavior of the chi-square technique depends on the uncertainty in the data, the 200 lidar profiles were created with averaging times ranging from 60s to 6000s. If N is the number of averaging times considered, the ensemble of lidar nitrogen signals will number $200\times N$ for each of the two simulated extinction profiles considered (figure 1) resulting in a total of $2\times200\times N$ simulated data profiles. From each of the simulated extinction profiles, there will also be $2\times200\times N$ extinction profiles obtained with each of the three polynomial fits: linear, quadratic and cubic. Finally the chi-square test will choose the most probable model for each altitude in each profile, resulting in $2\times200\times N$ profiles of extinction and corresponding uncertainty from the chosen model. For the simulations performed here, N was three since averaging times of 60, 600 and 6000 seconds were studied.

In figure 2 the frequencies with which the models are selected in the 200-profile ensemble

are displayed as a function of altitude and for different averaging times. For example in figure 2-c, corresponding to 600s averaging time, at the altitude of 2.5km the linear model (solid black line) was chosen very few times in the corresponding 200 extinction calculations while the quadratic (dotted line) and the cubic (dashed line) models were chosen with similar frequencies ($\sim$50%). In fact if we consider the points at 2.5 km in plot 2-b (60s averaging time) all of the three models are chosen a similar fraction of the time ($\sim$33%), while in plot 2-d (6000s averaging time) the linear model is never chosen and the quadratic and cubic models are chosen with frequency $\sim$50%. This means that at the fixed altitude of 2.5 km, the linear functional form describes less accurately the data with respect to the other two as we go from plot 2-b to 2-d. It is evident that the choice of model depends on the noise in the nitrogen signal. The results shown in plots 2-b to 2-d differ only in the averaging time and in particular plot 2-b corresponds to a larger uncertainty in the signal than plot 2-d. The uncertainty in the signal changes not only with the averaging time but also with altitude. In fact the points that correspond to higher altitudes in a lidar profile in general have larger uncertainty than those at lower altitudes. It is shown in any of the plots 2-b to 2-d that higher in the profiles the frequency of selection tends to be similar for the three models. Despite the fact that for any given profile within the ensemble of 200 there is a definite choice of the model, the variability of the data, given the noise level simulated, implies that all models are equally probable for the ensemble of simulations. This shows that in general as the uncertainty in the nitrogen signal increases the underlying functional behavior of the data is harder to reveal. This will be shown more clearly in figure 3.

An interesting result is that the linear model is never chosen as the most probable outside of the portions of profiles where the uncertainty is too large to allow for the underlying functional behavior to be revealed (i.e. where all three models are chosen approximately as often). An exception can be seen in the case of the lofted aerosol extinction layer case (Ext2) between 0 and 2.5 km, where the linear model has a peak in its likelihood. This implies that the choice of the model by means of the chi-square technique depends also on the shape of the extinction profile.

In figure 3 we investigate the model selection as a function of the signal uncertainty and the

extinction. In figure 3-a for example the white pixels indicate regions in which the frequency with which the linear model is chosen is 30% less than the cubic, grey pixels indicate when the linear and cubic frequencies differ by less than 30% and the black pixels correspond to lack of data. As noted already in figure 2, when the signal uncertainty is smaller the linear model is chosen less often than the other two. In figure 3, when the signal uncertainty is less than 0.2% the linear model is chosen more that 30% less than the cubic, but when the extinction is larger, for example 0.25 km^{-1}, even if the uncertainty in the signal is as much as 0.3% the probability of selection of the linear model is the same. Figure 3-b shows similar features. If we consider the line bisecting the diagrams, passing through the points (0,0) and (1.3,0.5), this line is dividing the diagram in two triangles. The triangle corresponding to large extinction and small signal uncertainty values is called upper triangle, the other is called lower triangle. The upper triangle in both the figures 3-a and 3-b (small signal uncertainty and large extinction) is the region in which the underlying functional behavior is more clearly distinguished.

To study the influence that the chi-square method technique has on the uncertainty in the calculated extinction, the average uncertainties for the 200 profiles ensemble and for a single averaging time (600s) are shown in figure 4. Both the average extinction uncertainty (4-a and 4-d) and the corresponding standard deviations (4-b and 4-e) are shown for the linear, quadratic and cubic models as well as for the model chosen by the chi-square technique (indicated by "chosen"). For the case of Ext1 (4-a), the mean extinction uncertainty corresponding to the linear model is up to a factor of four times larger than the uncertainty in the extinction obtained with either the quadratic or cubic fit below 2 km. In this altitude region the linear model was almost never chosen by the chi-square test (figure 2-b). For the case of Ext2 (4-d) the behavior is similar to the case of Ext1, with the exception of altitudes around 1km. In this range the linear model was considered as likely as the other two (2-g). These results show that there is on average a relationship between the likelihood of a model and the corresponding extinction uncertainty. In particular the linear model, when it is not the most probable, gives an uncertainty that is larger than the uncertainty calculated with the chosen model. This is shown also in figure 5. Six points along the extinction pro-

files are chosen, three for each simulated extinction profile (5a-b), and the frequency of the linear model having an uncertainty larger than the chosen model are shown as a function of the averaging time (5-c). For example in figure 5-c for averaging times equal to 60s, for the altitudes corresponding to the point indicated with A, the uncertainty in the extinction calculated with the linear model is 100% of the times (in the 200 points ensemble) larger than the extinction uncertainty calculated with the most probable model. In the same way corresponding to the altitude in 5-b indicated with D, the uncertainty calculated with the linear model is larger than the uncertainty with the chosen model 93% of the times. In the case of the 60s averaging time only in two cases, corresponding to the points indicated with C and F, the linear model uncertainty is smaller than the chosen model uncertainty less than 70% of the times. For larger averaging times that correspond to smaller uncertainties in the lidar signal the linear model gives an uncertainty larger than the chosen model more than 80% of the times. Therefore, this simulation shows that, as anticipated, the a priori choice of the linear model for the calculation of the aerosol extinction is generally unjustified and can have a strong influence in the estimate of the extinction uncertainty. In particular it results in larger uncertainties with respect to the most probable model chosen based on the chi-square test.

4. Application to experimental data

The chi-square technique was then applied to experimental data collected with the SRL[12] between 1996 and 1997. During that period, the SRL used a XeF laser, which has output spectrum centered at 351.1nm. The laser emits 400 pulses per second with an average power of 12-20W. The collection optics consists of a 0.76m Dall-Kirkham telescope. The scanning capability is made possible by the presence of a mirror that can rotate in a single axis allowing measurements along angles between 0 and $\sim$90 degrees from the vertical. The advantage of taking angle measurements is the possibility of retrieving data even below the minimum altitude defined in the vertical position by the overlap function. Information regarding the experimental configuration of the SRL during the time of these measurements may be found in[9]and.[10]

The data analyzed here are from two campaigns between 1996 and 1997 during which the scanning Raman lidar was deployed at the DOE/ARM site Southern Great Plains in Oklahoma. Even though the principal focus of these campaigns was the study and comparison of water vapor measurements from different instruments,[11] a large dataset of aerosol backscatter and extinction measurements was produced with the SRL. Time series of data at different angles were collected. A total of more than 80 hours of extinction profiles were analyzed using the chi-square technique. The data, corresponding to nighttime measurements, were selected for a percentage uncertainty in the resulting extinction smaller than 50%. This would allow rejecting portions of the data in which the extinction was either very small and/or the uncertainties very large.

With the purpose of investigating the relationship between the most probable model, the aerosol extinction and the signal uncertainty, all the data were combined to generate figure 6a which shows the most frequently chosen model as a function of the extinction and the signal uncertainty. The points in the dataset are divided in 7×7 cells (note how the cell size relates to both extinction and uncertainty) and among the points in one cell the most frequently chosen model is indicated. There is one cell that is labeled with Unclear which contains only two measurements among which the cubic model and the quadratic model were chosen an equal number of times by the chi-square technique. As with the simulated data, the linear model is chosen the least frequently in these experimental data. Moreover there is a general preference towards the cubic model ($\sim$ 70% likelihood).

As stated in the abstract, one of the goals of this research is to explore a general characterization of the extinction uncertainty as a function of the magnitude of aerosol extinction and the random uncertainty in the Raman lidar nitrogen signal. For such characterization, large datasets with measurements corresponding to different aerosol and atmospheric conditions are necessary. The dataset used in this paper corresponds to measurements obtained in three field campaigns during the same season (fall) and in the same geographic location, so they do not possess the characteristics necessary to draw general conclusions. Nonetheless, these results points toward the feasibility of generating such a general relationship and may be taken as a preliminary version of such a relationship. In figure 6b the average extinction

uncertainty is shown as a function of aerosol extinction and signal uncertainty. We can see that in the upper triangle, corresponding to higher extinction and smaller signal uncertainties, the average extinction uncertainty is always below 10%, while it is higher than 10% in the remaining region of the diagram but is never larger than 20%. Characterizations of the extinction uncertainty by means of these diagrams could allow evaluating the magnitude of the extinction uncertainty expected for different combinations of aerosol extinction and signal uncertainty. For example if a measurement of the aerosol extinction is expected to be on average 0.1 km^{-1} and the desired accuracy is less than 5%, according to the results of this study a signal uncertainty below 0.3% is necessary. At the same time if an accuracy of 10-20% is sufficient for the same value of the extinction, the uncertainty in the signal can be as large as 1.1%.

5. Implications

In the previous section we presented a general relationship between aerosol extinction and lidar extinction uncertainty. We use those results here to study the feasibility of accurately measuring the continental US background aerosol extinction. Also, the simulation executed here shows that on average the use of the chi-square technique to calculate aerosol extinction from Raman lidar data could reduce the uncertainty on the extinction with respect to the traditional technique (linear least square fit chosen *a priori*) by up to a factor of $\sim$4 (figure 4). This has an influence on the estimate of the uncertainties in the quantities that are used to evaluate the climatic effect of aerosols that depend on aerosol extinction and optical depth. Here a brief review of the potential influence of the reduction of the extinction uncertainty on some of these quantities, assuming that Raman lidar extinction measurements were used in each, is presented.

5.A. Continental US Background aerosol loading

Reference diagrams like the one in figure 6b can be useful for example to understand with which accuracy it is possible to measure the average background aerosol extinction. In 1988 a monitoring program was initiated called IMPROVE (Interagency Monitoring of Protected

Visual Environments)[13] with the objective to measure background visibility levels. One of the main results of this program was the calculation that the average background aerosol extinction over all the US territory was ~ 0.05 km^{-1} based on three years of measurements from stations distributed throughout the Unites States. If a similar background value was to be measured with the SRL, according to the diagram in figure 6b, it could be done with an uncertainty smaller than 10% if the uncertainty in the lidar signal was smaller than 0.3%. Using the configuration of the SRL as described, these measurements can be made during the nighttime with 300 seconds of averaging.

5.B. Aerosol direct effect

The aerosol direct effect is the effect that the aerosols have on the radiative balance of the atmosphere.[1] It is due both to aerosols absorbing and reflecting the radiation from the sun and the radiation emitted from the surface of the Earth. If an aerosol cloud is present a simple radiative transfer model accounting for surface emissivity, surface temperature, aerosol optical depth and aerosol temperature gives an equation for the aerosol radiative forcing. This is defined as the difference between the outgoing radiance at the top of the atmosphere when the aerosols are present and the same quantity in the absence of aerosols. A simple radiative transfer model that can be used to estimate the aerosol forcing F_{aer} (W m^{-2} sr^{-1} μm^{-1}) is:

$$F_{aer} = (1 - \varepsilon_{aer})\varepsilon_s B(\lambda, T_s) + \varepsilon_{aer} B(\lambda, \overline{T}_{aer}) - \varepsilon_s B(\lambda, T_s) \qquad (3)$$

where ε_{aer} is the emissivity of the aerosols, ε_s the emissivity of the surface, B is the black body emission function, T_s is the temperature of the surface and T_{aer} is the temperature of the aerosols. For example if the presence of aerosols in the atmosphere is as described from the aerosol extinction profile indicated with Ext1 in figure 1, we can estimate that T_{aer}=4° C, T_s=15° C, ε_s =0.95 and knowing that ε_{aer}=1-exp(-τ) , where τ is the optical depth of the aerosols, if one integrates the extinction profile to obtain optical depth, an extinction uncertainty like the one shown in figure 4a for the chi-square technique propagates

to an uncertainty in the aerosol forcing of ~0.8%. The resulting uncertainty in the aerosol forcing when the linear least square fit is used would be 3.2%. In this case using the chi-square technique as opposed to the traditional technique reduces the uncertainty on the aerosol forcing by a factor of 4. This means that the traditional technique would give a measurement of the aerosol forcing of -1.55±0.05 Wm^{-2} while the chi-square technique gives a measurement of -1.550±0.012 Wm^{-2}.

5.C. Aerosol indirect effect

As already stated, there is currently a high uncertainty on the knowledge of the effect that the aerosols have on cloud microphysics and lifetime (the aerosol indirect effect).[14] Recently a parameterization of the indirect effect was proposed by Feingold,[15] which defines the indirect effect as the ratio between the change in aerosol extinction α (or optical depth) and the change in cloud droplets effective radius r_e. The expression of the aerosol indirect effect is given by:[15]

$$IE = -\frac{\frac{\Delta r_e}{r_e}}{\frac{\Delta \alpha}{\alpha}} = -\frac{r_1 - r_2}{r_1 + r_2}\frac{\alpha_1 + \alpha_2}{\alpha_1 - \alpha_2} \tag{4}$$

where r_i are two different measurements of the effective radius and α_i are the two corresponding measurements of the extinction. The propagation formula for the uncertainty in IE, if only the uncertainties on the extinction (σ_1 and σ_2) are considered, is given by:

$$\sigma_{IE} = \sqrt{(2\alpha_1)^2 + (2\alpha_2)^2}\frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{(\alpha_1 + \alpha_2)^2} \tag{5}$$

If the uncertainty in extinction is 1% it results in an uncertainty in IE of 1.6%. If the uncertainty on the extinction is four times larger, as for example for the linear model with respect to the chi-square technique in the case of figure 4a below 1.5 km, the resulting uncertainty in the IE is ~6.5%. This means that a reduction of a factor of four in the uncertainty in the aerosol extinction results in a reduction of a factor of four in the uncertainty on IE.

5.D. CCN retrieval

The influence of aerosol on cloud microphysics starts with the influence on the concentration of Cloud Condensation Nuclei (CCN). An increase in the concentration of aerosols in the atmosphere causes an increase in CCN concentration that depends on the saturation conditions, since not all the aerosols are activated to grow to cloud droplets. For this reason the capability of measuring the concentration of CCN is essential to the improvement of the current knowledge about the cloud formation mechanisms. An algorithm for the calculation of the CCN concentration by extrapolation of measurements taken at the surface was introduced by Ghan.[16] The algorithm is based on the dependence of the CCN concentration on the aerosol extinction α(z) and its dependence on humidity f[RH(z)]. The expression of the CCN concentration at an altitude z is given by:

$$CCN(z) = CCN(z_0)\frac{f[RH(z_0)]}{f[RH(z)]}\frac{\alpha(z)}{\alpha(z_0)} \tag{6}$$

where z_0 is the ground altitude at which measurements of the CCN concentration are made directly (CCN(z_0)), f[RH(z_0)] and f[RH(z)] are the scattering dependence on humidity respectively at the ground and at the altitude z, that can be retrieved knowing the relative humidity profiles, and $\alpha(z_0)$ and α(z) are the extinction coefficients respectively at the ground level and at the altitude z. If we consider for example the case of Ext1 with $\alpha(z_0)$=0.5 km^{-1} and α(z)=0.17 km^{-1}, the uncertainties obtained with the chi-squared technique are respectively 0.0016 km^{-1} and 0.006 km^{-1}. If only the uncertainty σ_α on the extinction is considered, the uncertainty on the CCN concentration is found with the formula:

$$\sigma_{CCN} = \frac{\alpha(z)}{\alpha(z_0)}\sqrt{\left(\frac{\sigma_{\alpha(z)}}{\alpha(z)}\right)^2 + \left(\frac{\sigma_{\alpha(z_0)}}{\alpha(z_0)}\right)^2} \tag{7}$$

The uncertainties in the extinction result in an uncertainty in the CCN concentration equal to 3.5%. The uncertainties in the extinctions obtained with the linear model are respectively 0.037 $km^{-1}$and 0.015 km^{-1} and the resulting uncertainty in the CCN concentration is equal to 11%. In the case of the calculation of the CCN concentration the use of the chi-squared

technique as opposed to the linear fit would result in a reduction of the extinction uncertainty of a factor of 3.

5.E. Aerosol plume transport

Measurements of the aerosol extinction profile and the derived aerosol optical thickness are very useful also in case of transport of plumes from forest fires. An improvement of the ability to track aerosol clouds can lead to an improvement of our knowledge of the aerosol effects on climate. Given an aerosol cloud, from the measurement of its optical thickness and knowledge of the wind field it is possible to calculate the rate of mass transport. Infact the aerosol optical depth τ can be expressed as:[17]

$$\tau = \alpha B \tag{8}$$

Where α (m^2 g^{-1}) is the mass scattering coefficient of the aerosol and B (g m^{-2}) is the aerosol burden. If for example we assume that we have only sulfate aerosol in the plume, the mass scattering coefficient can be assumed to be equal to 5.0 m^2 g^{-1}.[18,19] Assuming an optical thickness of 0.5 and that the plume extends in the vertical plane for about 100 km^2 the volume density of the aerosol is ρ=10^{-4} g m^{-3}. If the wind velocity is 10 m s^{-1} the total mass transported per unit second will be then FM=100 kg s^{-1}. If the aerosol optical depth is measured with the Raman lidar technique and if we use σ_τ to represent the uncertainty in the aerosol optical depth then the uncertainty σ_{FM} on the flux of mass will be given by:

$$\sigma_{FM} = \frac{\sigma_\tau}{\tau} FM \tag{9}$$

Then it is clear that an increase of the uncertainty in the extinction of a factor of 4, which reflects on an equal increase in the uncertainty in the optical depth, would lead to an increase of a factor of four in the uncertainty in the mass flux.

6. Summary and Conclusions

An improved algorithm to calculate the aerosol extinction coefficient from a Raman lidar nitrogen signal has been validated. This algorithm uses the chi-squared test to choose the most probable least-square-fit model as opposed to choosing one *a priori*, which is the standard method of evaluating aerosol extinction from Raman lidar data. In order to correctly apply the chi-squared test the data need to follow a Gaussian distribution. To guarantee this, the traditional equation of the aerosol extinction must be reformulated.

The chi-squared test increases the confidence primarily on the accuracy of the aerosol extinction uncertainty. It was shown in a simulated set of measurements that in general the linear model is the least likely to fit the data. Consequently, based on the results of numerical simulations, the extinction uncertainty calculated with the chi-squared method was found to be on average a factor of 4 smaller than the uncertainty obtained with the linear model, while differences in the estimate of the extinction were smaller than 2.5%. All the data possessed a spatial resolution of 75m and were analyzed using a sliding window of five points for the regressions.

Additional simulations not presented in this paper indicated that the most probable model and its corresponding cumulative probability in general depend on the resolution of the data. The use of the chi-square test to distinguish among linear, quadratic and cubic models, as done here, is consistent with regression of 5 points. If more points need to be considered in the regression to improve the uncertainty in the retrievals, our tests indicate that, in general, re-binning the data so that approximately 5 points are used in the regression produces the cumulative probability closest to 0.5 for the least-squares fit when using the set of three models used here - linear, quadratic and cubic.

The ability of the chi-squared test to reveal the underlying functional behavior of the data, and therefore for this technique on average to yield smaller uncertainties than choosing a model *a priori*, depends on the uncertainty in the data (figure 3). In particular, based on the simulations shown here, for signal uncertainties smaller than 0.2% and aerosol extinction up to 0.4 km^{-1} the chi square test was generally able to determine the functional behavior

in a useful manner, whereas for signal uncertainties larger than 0.4% and aerosol extinction smaller than 0.2 km^{-1} the functional behavior was not clearly determined. In the experimental data shown here a Raman lidar signal uncertainty of 0.2% is achieved at a range of around 4 km for aerosol extinction values smaller than 0.1 km^{-1} using an averaging time of 300 seconds. Since many Raman lidar systems are capable of such measurements, the chi-squared technique can be generally useful for a more accurate estimate of tropospheric aerosol extinction and its uncertainty.

The chi-squared technique was applied to experimental data from two field campaigns indicating that the linear model was the least chosen by the chi-squared test and the cubic model was the most frequently chosen (70%). These results agree with the simulations and indicate further that the *a priori* selection of the linear model to calculate the aerosol extinction may be leading to inaccurate assessments of extinction uncertainty.

Finally, the extinction uncertainty resulting from the chi-squared technique has been shown in 2-D diagrams as a function of the aerosol extinction and the uncertainty in the Raman lidar nitrogen signal. For extinction values ranging from 0 to 0.26 km^{-1} and signal uncertainty ranging from 0 to 1.45% the extinction uncertainty is on average smaller than 30%. These diagrams can be considered preliminary "look-up tables" for estimating the aerosol extinction uncertainty given the value of aerosol extinction, which is accurately determined using an a priori selection of a linear model to regress the data, and Raman lidar measurement uncertainty. An example referring to the aerosol background averaged over the United States calculated during the monitoring program IMPROVE[13] of $\sim$0.05 km^{-1} was presented. If such an extinction is to be measured with an uncertainty less than 10% using Raman lidar, it is required that the uncertainties on the signal be smaller than 0.3%. On the other hand if the extinction uncertainty required is less than 20%, the uncertainty in the lidar signal can be as large as 1.1%. The conclusion is that the background aerosol loading in the continental US can be accurately measured at night with a 5 minutes average using Raman lidar systems with the characteristics of the SRL.

A reduction in the extinction uncertainty influences the uncertainty of quantities used to evaluate and study the effects of aerosols on the climate. The influence of errors in aerosol

extinction measurement has been studied based on the work presented here and assuming that a Raman lidar would provide the aerosol extinction data. A reduction of the extinction uncertainty by a factor of 4, like the magnitude of the difference between the linear model fit uncertainty and the chi-square technique results, causes a reduction by a factor of 4 in the aerosol radiative forcing uncertainty, a factor of 4 in the IE (a coefficient used to parametrize the aerosol Indirect Effect) uncertainty, a factor of 3 in the uncertainty in the CCN concentration retrieval and a factor of four in the uncertainty in mass flux estimate.

The results presented here are based on simulated retrievals of two extinction profiles and on nighttime Raman lidar measurements during the fall season at the same geographic location. Therefore they do not constitute a comprehensive database of either real or simulated Raman lidar aerosol extinction measurements under the full range of possible conditions. However, they nonetheless point strongly toward the conclusions 1) that the equation for aerosol extinction should be re-formulated prior to calculating extinction using any kind of regression technique due to the non-Gaussian statistics pertaining to the original form, and 2) that a model to fit the data should not be chosen *a piori* but rather determined based on the results of the chi-square test since the work done here indicated large increases in uncertainty estimates were generally obtained using a linear model but that the linear model was very unlikely to properly represent the data.

6.A. Aknowledgments

The work presented has been supported by the NASA Grant NNG04GC06A funded by the GSFC Earth Sciences Directorate for 2004-2006 and the NASA Radiation Sciences Program.

References

1. J.T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, C. A. Johnson, *Climate Change 2001: The Scientific Basis* (Cambridge University Press, Cambridge, UK and New York, 2001).
2. A. Ansmann, M. Riebesell, C. Weitkamp, "Measurements of atmospheric aerosol extinction profiles with a Raman Lidar," Opt. Lett., **15,** 746–748 (1990).
3. J. Bosenberg, V. Matthias., "EARLINET: A European Aerosol Research Lidar Network to Establish an Aerosol Climatology,"(2003) http://lidarb.dkrz.de/ec_earlinet/index.html.
4. G. Pappalardo, A. Amodeo, M. Pandolfi, U. Wandinger, A. Ansmann, J. Bosenberg, V. Matthias, V. Amiridis, F. De Tomasi, M. Frioud, M. Iarlori, Leonce Komguem, A. Papayannis, F. Rocadenbosch, X. Wang, "Aerosol lidar intercomparison in the framework of the EARLINET project. 3. Raman lidar algorithm for aerosol extinction, backscatter, and lidar ratio," Appl. Opt., **43**, 5370-5385, (2004).
5. D. N. Whiteman, "Application of statistical methods to the determination of slope in lidar data," Appl. Opt. **38**, 3360-3369 (1999)
6. G. Marsaglia, "Ratios of normal variables and ratios of sums of uniform variables," J. Amer. Stat. Assoc. **60**, 193-204, (1965)
7. P. R. Bevington, D.K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences 2nd Ed.* (McGraw-Hill, 1992)
8. D.N. Whiteman, G. Schwemmer, T. Berkoff, H. Plotkin, L. Ramos-Izquierdo, G. Pappalardo, "Performance modeling of an airborne Raman water vapor lidar,",App. Opt.,**40**, 375-390 (2001).
9. D. N. Whiteman, K. D. Evans, B. Demoz, D. O'C. Starr, E. W. Eloranta, D. Tobin, W. Feltz, G. J. Jedlovec, S. I. Gutman, G. K. Schwemmer, M. Cadirola, S. H. Melfi, and F. J. Schmidlin, "Raman lidar measurements of water vapor and cirrus clouds during the passage of Hurricane Bonnie," J. Geophys. Res., **106**, 5211-5225 (2001)
10. D. N. Whiteman, S.H. Melfi, "Cloud liquid water, mean droplet radius and number density measurements using a Raman lidar,", J. Geophys. Res., **104**, 31,411-31,419 (1999).
11. H.E. Revercomb, D.D. Turner, D.C. Tobin, R.O. Knuteson, W.F. Feltz, J. Barnard, J. Bosen-

berg, S. Clough, D. Cook, R. Ferrare, J. Goldsmith, S. Gutman, R. Halthore, B. Lesht, J. Liljegren, H. Linne, J. Michalsky, V. Morris, W. Porch, S. Richardson, B. Schmid, M. Splitt, T. Van Hove, E. Westwater, D. Whiteman, "The ARM program's water vapor intensive observation periods - Overview, initial accomplishments, and future challenges," Bull. of Am. Met. Soc., **84**, 217-236 (2003).

12. R. A. Ferrare, S. H. Melfi, D. N. Whieman, K. D. Evans, "Raman Lidar Measurements of Pinatubo Aerosols over Southeastern Kansas During November-December 1991," Geophys. Res. Lett. **19**, 1599-1602 (1992)
13. W. C. Malm, J. F. Sisler, D. Huffman, R. A. Eldred, T. A. Cahill, "Spatial and seasonal trends in particle concentration and optical extinction in the United States," J. Geophys. Res., **99**, 1347-1370 (1994)
14. S. Twomey, "The influence of pollution on the shortwave albedo of clouds," J. Atmos. Sci., **34**, 1149-1152 (1977)
15. G. Feingold, L. A. Remer, J. Ramaprasad, Y. J. Kaufman, "Analysis of smoke impact on clouds in Brazilian biomass burning regions: an extension of Twomey's approach," J. Geophys. Res., **106**, 22,907-22,922 (2001)
16. S. J. Ghan, D.R. Collins, "Use of in-situ data to test a Raman Lidar-based Cloud Condensation Nuclei remote sensing method," J. Atmos. Oc. Tech., **21**, 387-394 (2004)
17. J. Lelieveld, J. Heintzenberg, "Sulfate Cooling Effect on climate through in-cloud oxidation of anthropogenic SO2," Science, 258, 117-120 (1992)
18. R.J. Charlson, S.E. Schwartz, J.M. Hales, R.D. Cess, J.A. Coakley, J.E. Hansen, D.J. Hoffman, "Climate forcing by anthropogenic aerosol," Science, 255, 423-430 (1992)
19. R.M. Hoff, S.P. Palm, J. A. Engel-Cox, J. Spinhirne, "GLAS long term transport observation of the 2003 California forest fire plumes to the Northeastern US", Geoph. Res. Let., 22, DOI 10.1029/2005GL023723, L22S08

Table 1. Configuration Of SRL during the campaigns VWIOP, WVIOP2 and VWIOP3.

Laser Wavelength	351.1 nm
Raman wavelength	382.4 nm
Energy Per Pulse	40 mJ
Primary Telescope	0.76 m
Secondary Telescope	0.13 m
Telescope f number	5
FOV	0.002 rad
Bin time	0.5 μs
Laser Rep Rate	400

List of Figure Captions

Fig.1. Simulated extinction profiles that were used to simulate nitrogen signals. The one labeled Ext1 represents a typical aerosol content decreasing with altitude in the boundary layer. The profile labeled Ext2 represents a lifted aerosol cloud with a raising portion between 0 and 1.5 km that could simulate for example aerosol hygroscopic growth.

Fig.2. Panels showing the frequencies with which each model is chosen throughout the 200 profiles ensemble. Panels (a) and (e) show the simulated extinction profiles corresponding to the results that are shown respectively in panels (b)-(d) and in panels (f)-(h) for averaging times 60s, 600s and 6000s.

Fig.3. Diagrams showing for both simulated extinction profiles Ext1 (a) and Ext2 (b) showing the ranges of extinction and signal uncertainty for which the difference in frequencies between linear and cubic model are larger than 30% (white pixels). The black pixels indicate that there are no data available.

Fig.4. Panels showing the average uncertainties on the extinction calculated with the linear model, quadratic model and cubic model only together with the model chosen by the chi-square test (chosen) for 600s averaging times for both the simulated extinction profiles. Panels (a) and (d) show the extinction uncertainty profiles averaged over the 200 points, panels (b) and (e) show the standard deviations of each average profile and (c) and (f) show the corresponding simulated extinction.

Fig.5. Panels showing the six points selected for illustrative purposes (a,b). The plot (c) shows that in general more than 80% of the times the linear model results in extinction larger than the chi-square technique.

Fig.6. Diagrams showing the results of the experimental dataset. Diagram (a) shows the

most frequently chosen model as a function of extinction and signal uncertainty, while diagram (b) shows in a similar diagram the average extinction uncertainty.

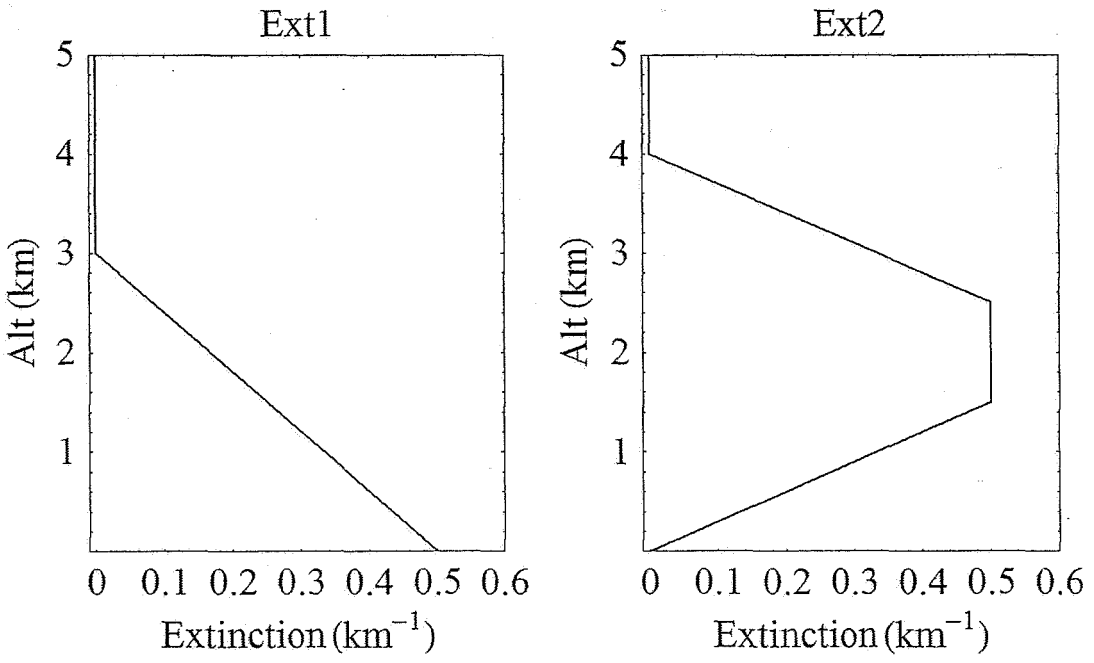

Fig. 1. Simulated extinction profiles that were used to simulate nitrogen signals. The one labeled Ext1 represents a typical aerosol content decreasing with altitude in the boundary layer. The profile labeled Ext2 represents a lifted aerosol cloud with a raising portion between 0 and 1.5 km that could simulate for example aerosol hygroscopic growth.

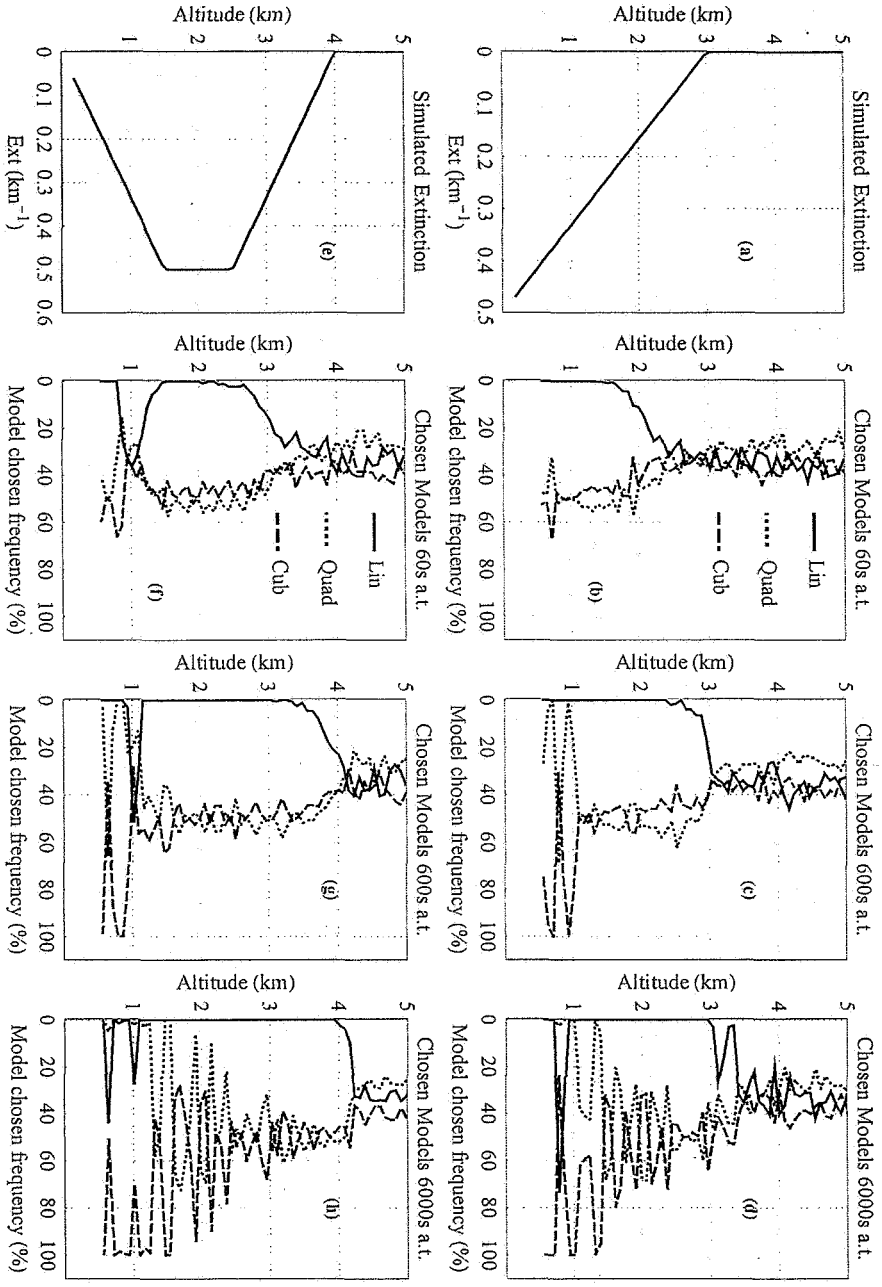

Fig. 2. Panels showing the frequencies with which each model is chosen throughout the 200 profiles ensemble. Panels (a) and (e) show the simulated extinction profiles corresponding to the results that are shown respectively in panels (b)-(d) and in panels (f)-(h) for averaging times 60s, 600s and 6000s.

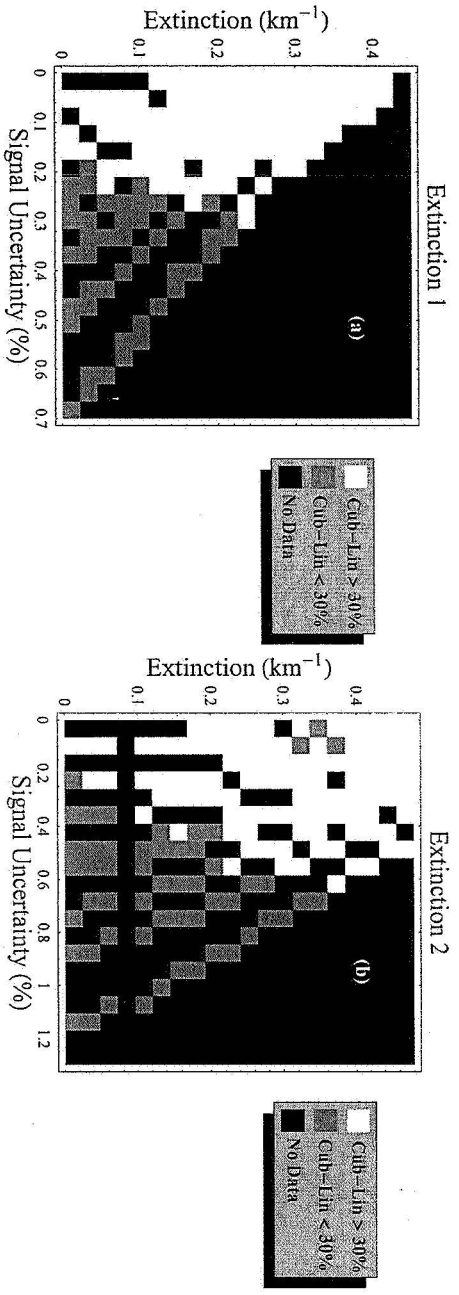

Fig. 3. Diagrams showing for both simulated extinction profiles Ext1 (a) and Ext2 (b) showing the ranges of extinction and signal uncertainty for which the difference in frequencies between linear and cubic model are larger than 30% (white pixels). The black pixels indicate that there are no data available.

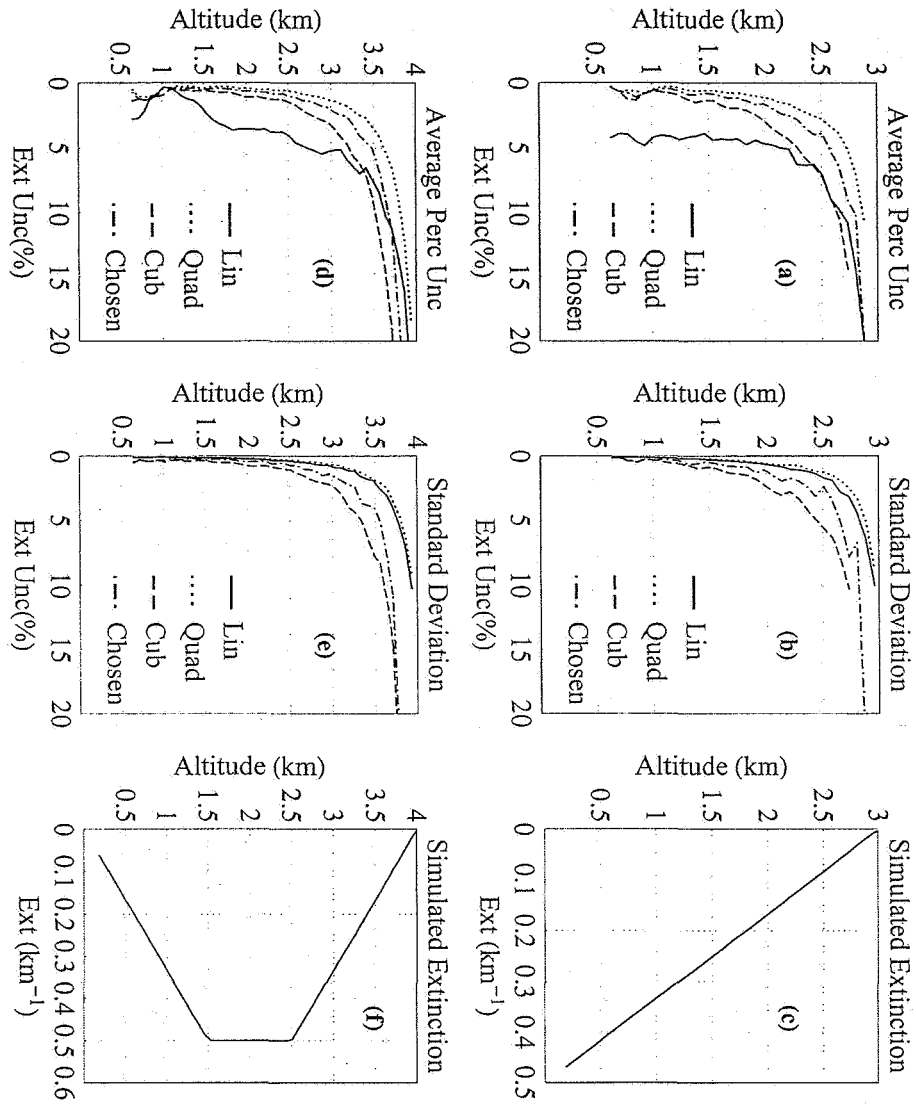

Fig. 4. Panels showing the average uncertainties on the extinction calculated with the linear model, quadratic model and cubic model only together with the model chosen by the chi-square test (chosen) for 600s averaging times for both the simulated extinction profiles. Panels (a) and (d) show the extinction uncertainty profiles averaged over the 200 points, panels (b) and (e) show the standard deviations of each average profile and (c) and (f) show the corresponding simulated extinction.

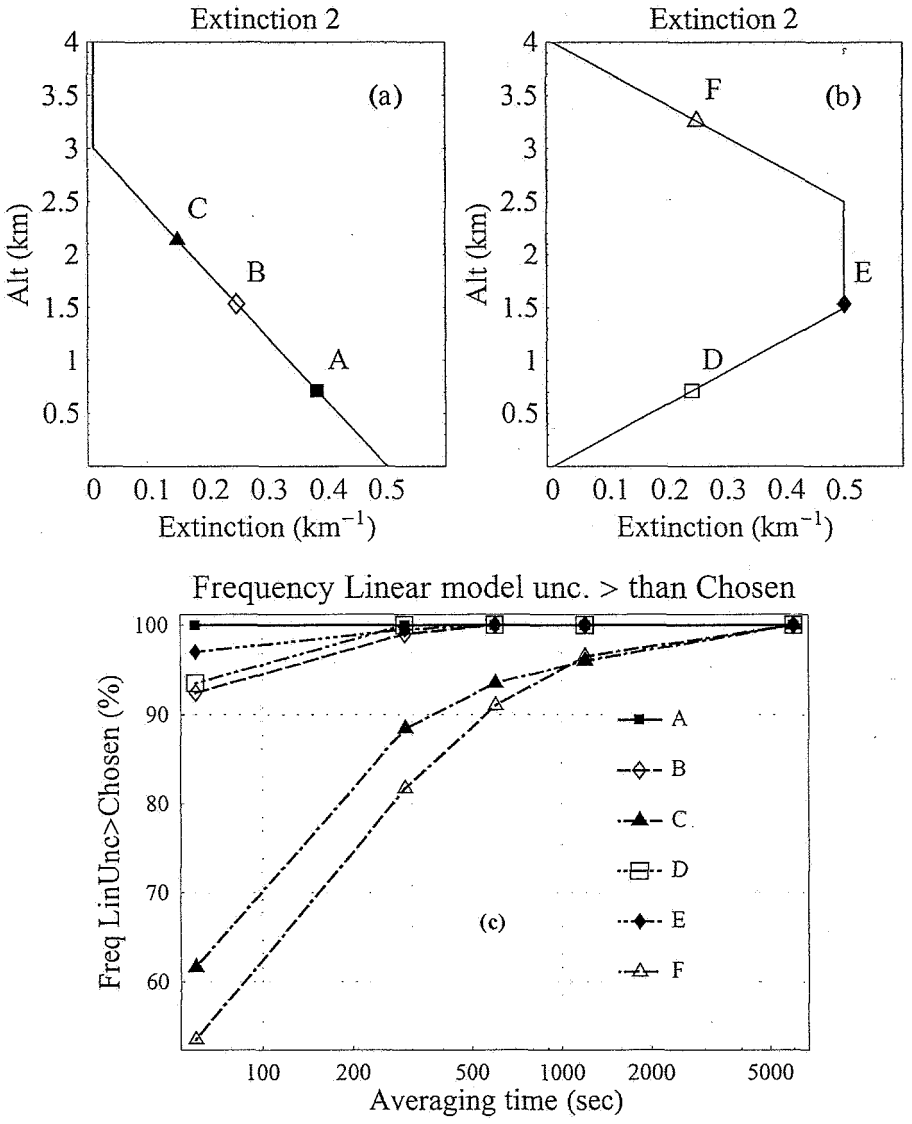

Fig. 5. Panels showing the six points selected for illustrative purposes (a,b). The plot (c) shows that in general more than 80% of the times the linear model results in extinction larger than the chi-square technique.

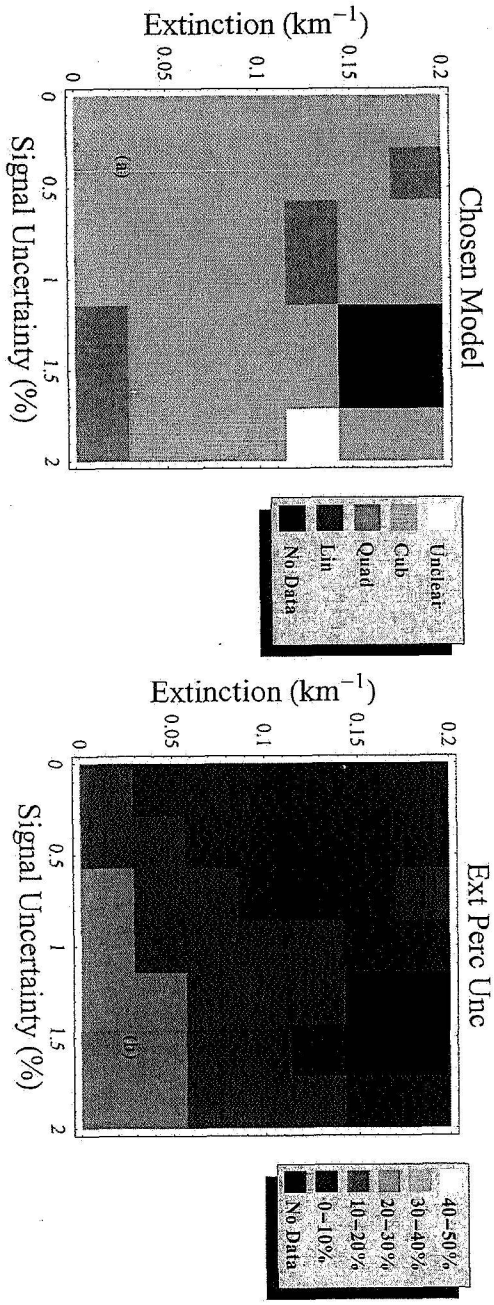

Fig. 6. Diagrams showing the results of the experimental dataset. Diagram (a) shows the most frequently chosen model as a function of extinction and signal uncertainty, while diagram (b) shows in a similar map the average extinction uncertainty.